AHRQ Grant Final Report

Title of Project:

Exploring Clinically-relevant Image Retrieval for Diabetic Retinopathy Diagnosis

Principal Investigator and Team Members:

Li, Baoxin, Ph.D., M.S., Principal Investigator

Ragav Venkatesan, M.S., Graduate Research Assistant

Parag Chandakkar, M.S., Graduate Research Assistant

Helen K. Li, M.D., Consulting Ophthalmologist

Grace Chen, Ph.D., Consulting Statistician

Organization:

Arizona State University-Tempe Campus

Inclusive Dates of Project:

August 1, 2011 – May 31, 2014

Federal Program Officer:

Steve Bernstein

Acknowledgement of Agency Support

Agency for Healthcare Research and Quality

Grant Award Number:

R21 HS 019792

Abstract

Purpose: The objectives of the project were to develop a content-based retrieval system for referencing diabetic retinal images to improve diagnosis and to develop a prototypical DR image management system to improve reviewers' diagnostic performance.

Scope: Diabetic retinopathy (DR) is a common cause of blindness. Despite advances in diabetes care, visual impairment is still a devastating complication. The project intended to develop new methodologies and computer software tools for assisting clinicians in reviewing fundus images for DR evaluation.

Methods: The key approach of the project was to develop a machine-learning-based image retrieval engine that would be able to retrieve a number of reference images from a database of standard images or previously-diagnosed images, for any given new fundus image to be diagnosed. A key aspect of the method was on designing feature extractors that can detect image features appropriate for distinguishing DR lesions. The method was evaluated by examining whether the retrieved images have similar DR severity to that of the new image.

Results: A database containing image samples of two representative DR lesions, neovascularization and microaneurysm, was used to evaluate the developed retrieval engine. On average, the proposed method was able to outperform other approaches in the literature. A retrieval interface was also developed to support the deployment of the retrieval engine in a user-friendly fashion. The source code and the dataset were posted on a Website for other researchers to further explore.

Keywords: Diabetic retinopathy, fundus image, image retrieval, computer-aided diagnosis, machine learning.

Purpose

Diabetic retinopathy (DR) is a common cause of blindness in working populations [1]. Despite advances in diabetes care over the years, visual impairment is still a potentially devastating complication. Studies have shown that timely DR diagnosis and treatment can significantly reduce the risk of severe vision loss [2]. Maximizing the efficiency and accuracy of assessing DR severity level could help prevent vision disabilities and their resulting high cost to the society. In practice, different eye care providers have different abilities in assessing diabetic retinopathy, and their accuracy in evaluating the stages of the disease depends upon training and experience. This remains to be the case in diagnosis based on digital retinal imaging, which has quickly become an alternative to traditional face-to-face evaluations. Although evaluations are being made from digital retinal images, the process is laborious and prone to error or reviewer fatigue. For this reason, many researchers have been exploring automated detection and evaluation of diabetic retinal lesions (e.g., [3-16]). Potential benefits of automated DR diagnosis include improved consistency and speed over human reviewers. However, no current computer-based systems for evaluating DR incorporate the kinds of experience and all of the variables that clinicians cognitively apply in detecting and assessing the severity of DR, and thus much more efforts are still needed to improve the performance of such systems. This project explored a new methodology for retrieving clinically-relevant images from archived database with expert annotations (diagnosis information), for a given novel image, which can provide a reviewer with *instant* reference to annotated images from a database.

The ultimate goal was to decrease blindness in diabetic patients. In particular, this research project explored an innovative method to retrieve clinically-relevant images for facilitating timely and accurate evaluation of DR. Images are considered as being clinically relevant if they contain the same types of lesions with similar severity levels. The investigators hypothesized that the computer-based method of referencing diagnostic retinal images to a library of images representing ranges of lesions and severity levels would significantly improve a clinician's diagnosis, if the referencing is based on clinical relevance. The project had the following specific aims:

Specific Aim 1. **Develop a content-based retrieval system for referencing diabetic retinal images to improve diagnosis**. The investigators utilized machine-learning techniques to develop retrieval algorithms that are content-based (data-driven, no metadata assumed) and are capable of determining clinical relevance. The method is significantly different from existing approaches to computer-based evaluation of DR that typically attempt to explicitly code human knowledge into a computer algorithm, which remains a challenging task and hence the performance bottleneck of those approaches.

Specific Aim 2. Develop a prototypical DR image management system to improve reviewers' diagnostic performance. In building the system by incorporating the content-based retrieval technique, the team aimed at addressing key issues such as data structures for knowledge representation that supports efficient retrieval, interface schemes that supports intuitive annotation, and visualization schemes (for presenting the retrieved results) to improve a clinician's performance in accuracy and speed of diagnosis.

The research targeted at providing an innovative way of exploiting vast expert knowledge hidden in libraries of digital DR images. The method that retrieves only clinically-relevant images renders it possible for a clinician to efficiently utilize the expert knowledge that is otherwise hidden in a database. Our retrieval method is unique in that it focuses on clinical relevance of the retrieved images instead of their similarity in terms of low-level image features to a query image (the image to be diagnosed).

Scope

Accordingly to the estimate in an earlier report [17], blindness caused by diabetes costs the United States more than $500 million annually. A World Health Organization collaborative study projected the global diabetic burden to affect 221 million people by 2010 [18]. In August 2010, a study on this subject [19] reported alarming rates of 28.5% and 4.4%, respectively, for diabetic retinopathy (DR) and vision-threatening DR among US adults with diabetes. Hence, with the increased prevalence of diabetes in the US, it is believed that the estimated cost due to blindness caused by diabetes has become much larger at present. Research resulting in better ways of evaluating diabetic retinopathy (DR) should significantly improve diagnosis of the disease and a reduction of its health care costs. Computer-based methods could potentially enhance the speed and accuracy of large-scale evaluation/screening programs and thus significant efforts have been invested in this field. One good example of such methods is automated lesion detection in digital fundus images. Unfortunately, to date there is no automated system that can perform DR lesion detection with the accuracy that is comparable to a human expert. Further, a task that is more difficult for ophthalmologists – fast and accurate severity grading of DR – has not been adequately addressed by existing systems, and an ophthalmologist's painstaking visual examination of a digital retinal image and mental or physical comparison with standard images are still the ultimate method for identifying and assessing DR. The proposed research aimed at filing this gap in both technical capability and clinical practice by developing a computer-based system with the innovative idea of content-based retrieval of DR images with clinical relevance.

The goal was to develop a novel technology to retrieve and reference images based on their contents that are clinically relevant to a novel image to be diagnosed. Therefore, the retrieval system can provide an ophthalmologist with both standard sample images and previously-diagnosed images of similar lesions and severity, and thus effectively enables knowledge sharing/reutilization among experts. Such a knowledge-sharing scheme can be especially useful for inexperienced doctors and medical students in training, as they can gain instant access to expert knowledge through the automatically retrieved reference images. It may also be a key technique for mobile DR screening clinics that are not staffed by experienced doctors. The retrieved images may also include past images with confirmed diagnosis by the same clinician, thus enabling instant access to his/her past cases. Additionally, the development of a DR image database may facilitate the integration of retinal image databases with other hospital information systems for creating management guidelines and determining appropriate treatment. This is nearly impossible in current practice where the annotations (diagnosis information) are often kept on separate paper charts and thus are hardly usable by a database system.

Methods

Overall Research Design

Studies have shown that timely DR diagnosis and treatment can significantly reduce the risk of severe vision loss. Thus maximizing the efficiency and accuracy of detecting DR and assessing its severity level can improve the practice on preventing vision disabilities due to DR, leading to reduction of the high health cost associated with DR and its complications. Automating the assessment of DR images not only has the potential of improving the efficiency and accuracy of the practicing physicians, but also may bring about automated pre-screening systems that can be deployed in settings/locations where experienced ophthalmologists may not be available for timely detection of DR. For this reason, recent years have witnessed fast developments of computer-aided systems/approaches for analysis of DR images, including the studies on automated detection and evaluation of diabetic retinal lesions. However, no current computer-based systems for evaluating DR incorporate the kinds of experience and all of the variables that clinicians cognitively apply in detecting and assessing the severity of DR, and thus much more efforts are still needed to improve the performance of such systems. This project developed a new concept in designing a computer-aided DR image analysis system: for a given image to be diagnosed, the system retrieves, from a database of previously-diagnosed images, *clinically-relevant* images such that the pre-stored expert annotations (diagnosis information) of these images become instantly available for reference. This represents a new means of diagnosing DR that exploits the knowledge- and information-laden archival database. Furthermore, the development of the novel image retrieval system also naturally resulted in automated DR evaluation algorithms with potentially improved performance compared with existing methods, since the focus of our system is on clinical relevance rather than low-level image features. This concept is schematically illustrated in Figure 1.

To achieve the goal manifested in the above concept, this research project proposed to develop an innovative, computational method to retrieve clinically-relevant images for facilitating timely and accurate evaluation of DR. Images are considered as being clinically relevant if they contain the same types of lesions with similar severity levels. The investigators hypothesized that the computer-based method of referencing diagnostic retinal images to a library of images representing ranges of lesions and severity will significantly improve a clinician's diagnosis, if the referencing is based on clinical relevance. In its actual implementation, the approach may also be used to build a system that can triage incoming images into those in need of a physician's more attention and those that can be automatically labeled.

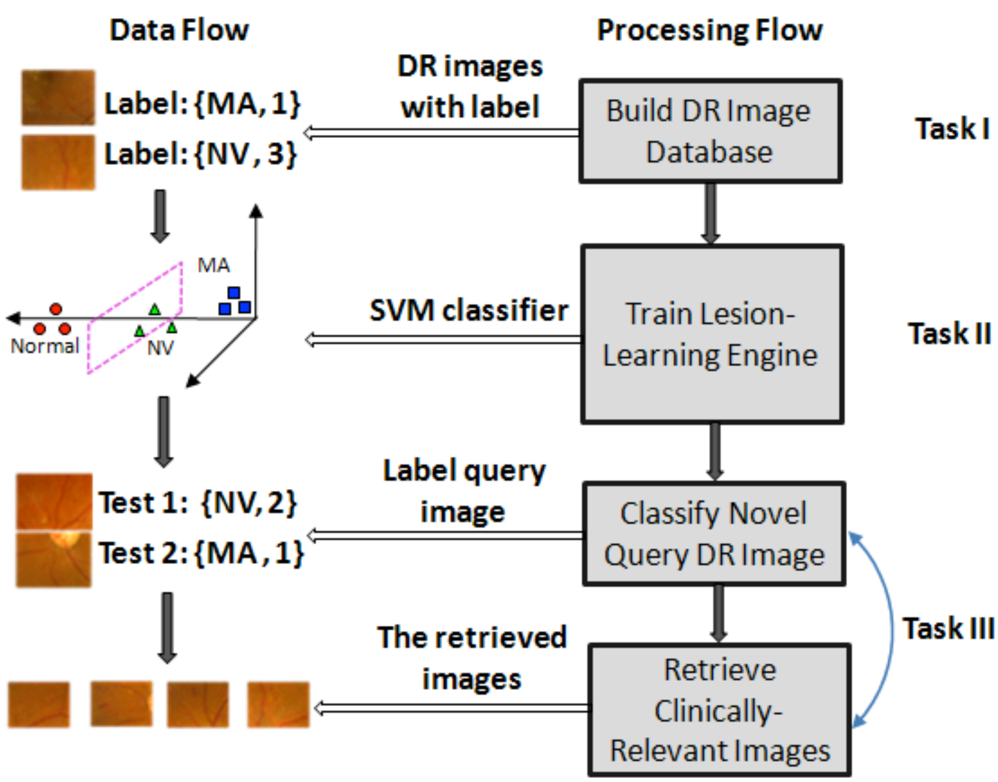

Figure 1. The general framework of the proposed system and its operation and components.

Key Technical Components

In realizing the concepts manifested in the overall design of Figure 1, the project was focused on developing several technical components, which are elaborated below.

Designing Core Algorithms for Building Automatic Lesion-Learning Engines

In this task, we used labeled data to obtain two kinds of classifiers: lesion-type classifier and lesion-severity classifier. The lesion-type classifier classifies an image according to the types of lesions it may contain. In general, the types of DR lesions include microaneurysm, retinal hemorrhages, hard exudates, intra-retinal microvascular abnormalities, and neovascularization, etc. In our initial exploration of the proposed research, the lesion-type classifier classifies an image into one of the three categories: microaneurysm (MA), neovascularization (NV) and normal (no lesion is present). MA is an important lesion to study because it is a telltale lesion of diabetic retinopathy. Since it is the earliest clinically recognizable disease lesion, it is a sign of early diabetic retinopathy. MA is also a surrogate measure of diabetic retinopathy severity. On the other hand, NV signifies a severe DR stage. Hence in a sense MA and NV are the most important lesion types and accordingly we choose to focus on them initially in order to maximize the outcome of the research given limited time and resources. Some images are likely to have multiple types of lesions. For example, an image with NV is very likely to have MA as well. In such a case, we label the image to the category which corresponds to the advanced phase of the lesion development, e.g. an image containing both NV and MA will be labeled as NV.

We proposed to use McMIL to train the classifiers and classify novel query images. In this setting, let D be the labeled data which consists of a set of m images $B=\{B_1, B_2, \ldots, B_m\}$ and their labels $L=\{l_1, l_2, \ldots, l_m\}$, i.e. D is augmented with the set of labels into $D=\{<B_1, l_1>, \ldots, <B_m, l_m>\}$. The images in the training set are collected from several sources. The label l_i is a vector $l_i = \{t_i, s_i\}$ giving us the labels in terms of both lesion type and severity level. t_i can take three values 1, 2 or 3, corresponding to MA, NV and normal classes. s_i can take three values 1, 2, or 3, indicating three severity grades: light, moderate, and proliferative. We use pairs of (B_i, t_i), $I = 1,\ldots, |\boldsymbol{D}|$ to train the lesion-type classifier, and we use pairs of (B_i, s_i), $I = 1, \ldots, |\boldsymbol{D}|$ to train the lesion-severity classifier.

Developing Algorithms for Clinically-Relevant Image Retrieval

In this task, we developed algorithms to retrieve clinically-relevant images from a database, for a given query image. The processing flow of this step is illustrated in Figure 2. It includes a 2-step process: classification and retrieval, as described in the following.

1. Classification**:** The classification of a query image proceeds as the following two steps: In the first step, we use the decision function of the *lesion-type* classifier to obtain the lesion-type label, i.e. one of the three classes {MA, NV, normal}. In the second step, knowing the lesion-type label of the query image, we use the decision function of the *lesion-severity* classifier corresponding to that lesion-type class to obtain the lesion severity label. After these two steps, a query image is associated with two labels, e.g. {MA, 1}, one indicating the lesion-type, and the other one denoting the severity level of that type of lesion. Note that there is no severity classification if a query image is classified as "normal". In the scope of this project, due to the lack of elaborate labels for the severity, the study was exclusively focused on the lesion types only, without further elaborating their severity levels.

2. Retrieval: The retrieval stage involves two steps. In the first step, the system first selects the images from the database whose lesion-type and lesion-severity label are the same with the two labels of the query image; it then compares the *lesion-type* feature vector of the query image with the *lesion-type* feature vector of each of the selected images; finally it retrieves the top K_1 nearest neighbors in terms of the Euclidean distance in the *lesion-type* bag feature space. Although not done in this project, conceptually, it is possible to further retrieve K_2 images from the K_1 whose *lesion-severity* features are similar to the *lesion-severity* features of the query image, which is achieved by choosing the top K_2 nearest neighbors from the K_1 images in terms of the Euclidean distance in the *lesion-severity* bag feature space. The K_2 images are clinically-relevant to the query image in the following senses: (*i*) they share the same lesion-type and lesion-severity label, and (*ii*) their visual appearances in terms of both lesion-type and lesion-severity are similar, these include color, texture, shape of lesion, etc. At this point, it is worth emphasizing again that most existing work on content-based retrieval can obtain images that are similar to the query image only in the second sense.

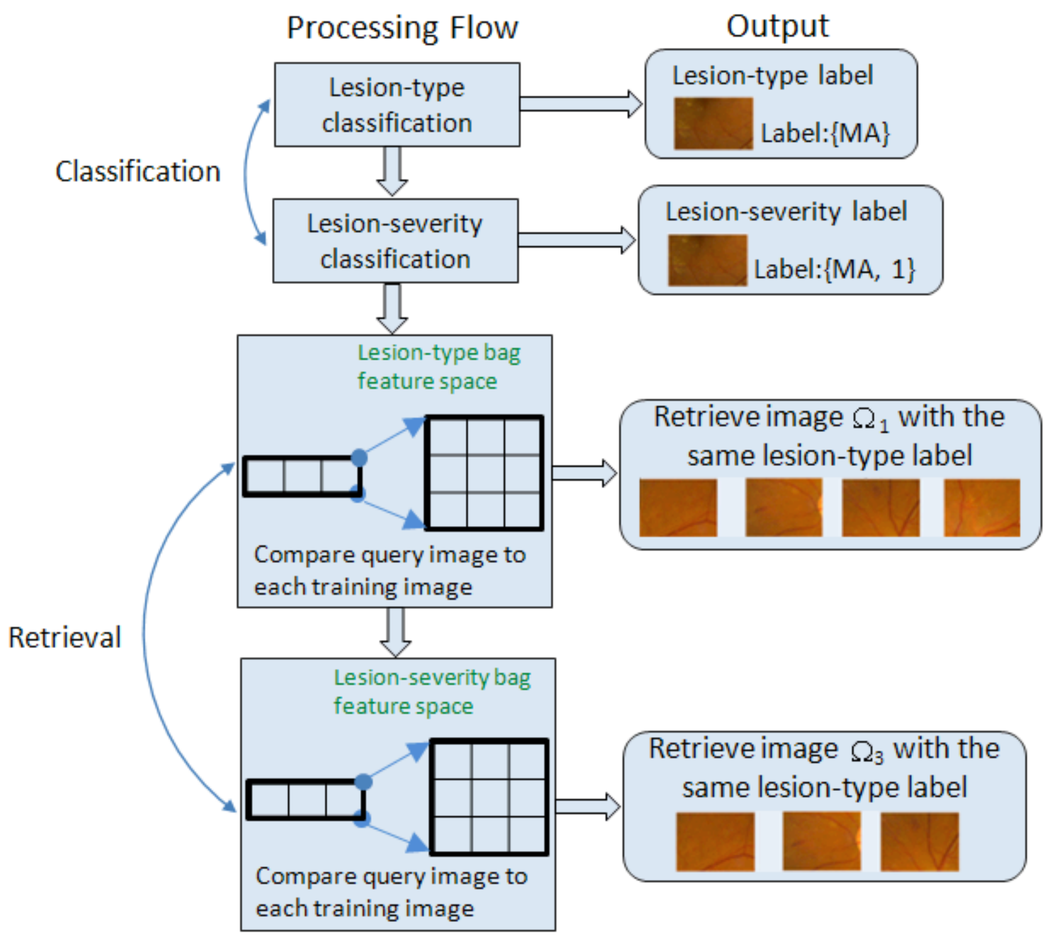

Figure 2. The processing flow of clinically-relevant retrieval.

Designing DR-image-specific Image Features

Color is a distinguishing characteristic for DR images. However, accurate modeling of the color spectrum of DR images can be challenging due to the variation in lighting in the images of same class. This is illustrated in Figure 3. In spite of belonging to the same class, the images in the first and last row have a large difference in their color spectrum. On the other hand, images belonging to different classes may have very small difference in their color spectrum and often that difference is localized. Figure 4 illustrates this perfectly.

We use color correlogram to represent DR images as it encodes global distribution of local spatial correlation of colors. We modify it to be invariant to the lighting changes and shifts. It is also tuned to the peculiar spectrum of DR images. Our key approach was to design a spectrally-tuned Color Correlogram, which takes into consideration the special color distributions of typical DR images. The details of this technique have been reported in our publications (listed in the section "List of Publications and Products").

Figure 3. Fundus image samples: normal (top row), NPDF (middle row), and PDR (bottom row).

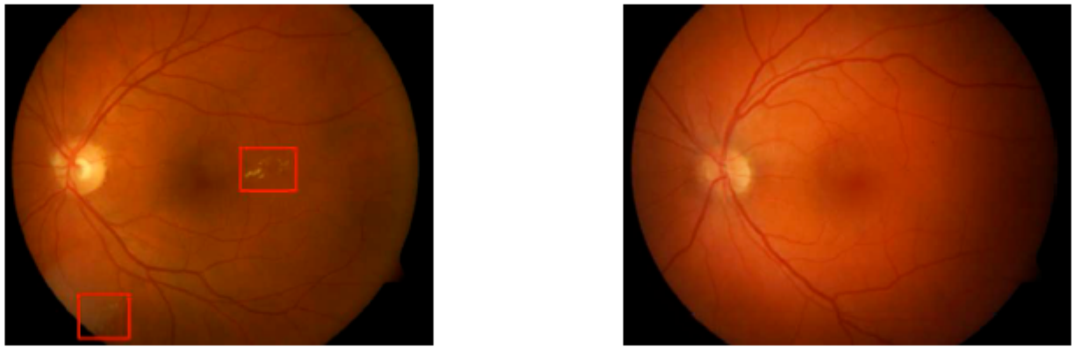

Figure 4. An NPDF image (left) and a normal image (right).

Results

In the following, the results of the project are summarized based on how they contribute to addressing the Specific Aims of the project.

Specific Aim 1. **Develop a content-based retrieval system for referencing diabetic retinal images to improve diagnosis**.

Towards addressing this Specific Aim, two major technical accomplishments were made as the results of the project efforts. These are summarized below. The detailed presentations of these results were reported in publications/manuscripts to be listed in the next section "List of Publications and Products".

1. Classification of Diabetic Retinopathy Images Using Multi-Class Multiple-Instance Learning Based on Color Correlogram Features.

In this effort, we proposed a spectrally-tuned AutoCC feature and an MIL framework based on this feature for DR image classification. Experiments with comparison with commonly-used image classification approaches demonstrated that the proposed method is able to achieve significant performance gain. Following experiment highlights the significant results of this effort.

The proposed features were evaluated against some commonly-used features in the medical imaging literature, such as Gabor features [20] and semantic of neighborhood color moment histogram features (HNM) [21]. These features are used in conjunction with the powerful SVM classifier to form classification schemes. Furthermore, we also considered a state-of-the-art classification scheme in computer vision, SIFT+BoW+SVM, which has been found effective in many computer vision tasks. (Whenever SVM is used, our implementation was based on the well-known LibSVM toolbox [22].) Finally, to provide a case that purely evaluates the benefit of the proposed spectrally-turned AutoCC feature, we also evaluated a scheme using the original AutoCC features in [23] together with the MIL approach. All these approaches are listed in the first column of Table 1.

Table 1. Mean accuracy of various methods. The dataset consists of 425 images. For details, see [Chandakkar et al 2011]

Approach	Mean Accuracy
SIFT+BoW+SVM	51.14 %
Gabor features+SVM	64.71 %
HNM + SVM	75.76 %
Original AutoCC+MIL	78.01 %
Proposed Algorithm	87.61 %

2. Retrieving clinically relevant diabetic retinopathy images using a multi-class multiple-instance framework

In this effort, we developed a novel approach for retrieving clinically-relevant DR images. The approach consists of a feature space which is spectrally-tuned to the DR spectrum. Feature space makes optimal utilization of the quantization scheme and thus produces a better representation of the image even with less number of bins. It makes sure that shades in DR images are almost uniformly spread across all the quantization bins, thereby creating a feature space with much higher entropy. Approach also consists of a multi-class multiple-instance retrieval framework called MIRank-KNN which uses minimal Hausdorff distance. The results using the proposed approach are reported bellow, with comparison with other state-of-art retrieval frameworks in the literature.

We first present four evaluation metrics: (1) ≥k hit-rate; (2) mean accuracy at k-th rank; (3) precision at k-th rank (P@k); and (4) mean average precision (MAP). The ≥k hit-rate (HR) is defined as the percentage of images for which at least k relevant images were retrieved. Mean accuracy at k-th rank denotes the percentage of relevant images retrieved at that particular rank. Precision at k-th rank measures the relevance of top k images in the ranking result with respect to the query image. We average P@k values for all the queries to get a single P@k value.

We compared our approach with two state-of-art image retrieval systems: textured image retrieval using Gabor features [20], medical image retrieval using HNM [21]. Table 2 is for the mean accuracy and P@k metric, and Table 3 for ≥k hit-rate.

Table 2.
MEAN ACCURACY AND PRECISION AT k^{th} RANK (IN %)

	$Acc@1/$ $P@1$	$Acc@2/$ $P@2$	$Acc@3/$ $P@3$	$Acc@4/$ $P@4$	$Acc@5/$ $P@5$
Gabor	73.43/73.43	72.62/73.02	68.76/71.60	67.75/70.64	70.39/70.59
HNM	77.69/77.69	74.04/75.86	72.01/74.58	73.83/74.39	69.57/73.43
Proposed	84.38/84.38	81.54/82.96	83.37/83.10	79.51/82.20	80.93/81.95

Table 3.
$\geq k$ HIT-RATE (in %)

	≥ 1 HR	≥ 2 HR	≥ 3 HR	≥ 4 HR	≥ 5 HR	Mean Acc.	MAP
Gabor	93.91	86.61	76.06	60.24	36.10	70.59	79.68
HNM	94.52	87.22	77.28	64.50	43.61	73.43	82.49
Proposed	96.55	91.48	87.42	76.88	57.40	81.95	87.6

A mean confusion matrix is created to better understand the results of our approach. It is shown in Table 4. The following example illustrates the process of constructing a confusion matrix. Suppose all images are queried one-by-one resulting in n retrieved images. The first row of the confusion matrix shows that 81.59% of the n images were normal, when a normal image was queried. Similarly, 13.78% images were NPDR and 4.63% images obtained were PDR, while querying a normal image. The second and third row can be similarly explained.

Table 4.

MEAN CONFUSION MATRIX (IN %)

	Normal	NPDR	PDR
Normal	81.59	13.78	4.63
NPDR	9.81	80.50	9.69
PDR	8.45	7.86	83.69

Specific Aim 2. Develop a prototypical DR image management system to improve reviewers' diagnostic performance.

Partially in accommodating the tasks performed in addressing Specific Aim 1, we already designed proper data structure for knowledge representation that supports efficient retrieval. In integrating the retrieval engine and the classifier into a DR image management system, we developed a software tool/system, whose interface is illustrated below (Figure 5). As can be seen from this interface, which maintains a DR database in the background, one can load a new picture (left panel, top), and then make a query to the database. Top matches from the dataset, as obtained by the retrieval engine, are then displayed on the screen (middle panel, the column of five images). The user may choose any of the retrieved images for further inspection (right panel). If the database images have any prior annotations (such as diagnosis results), such information may be displayed on the screen as well, so that the user may use that as reference in evaluating the new image.

We note that the above serves only as a highlight/summary of the results from the project, considering the page limit. The detailed results and analysis have been documented in the publications/manuscripts listed under "List of Publications and Products".

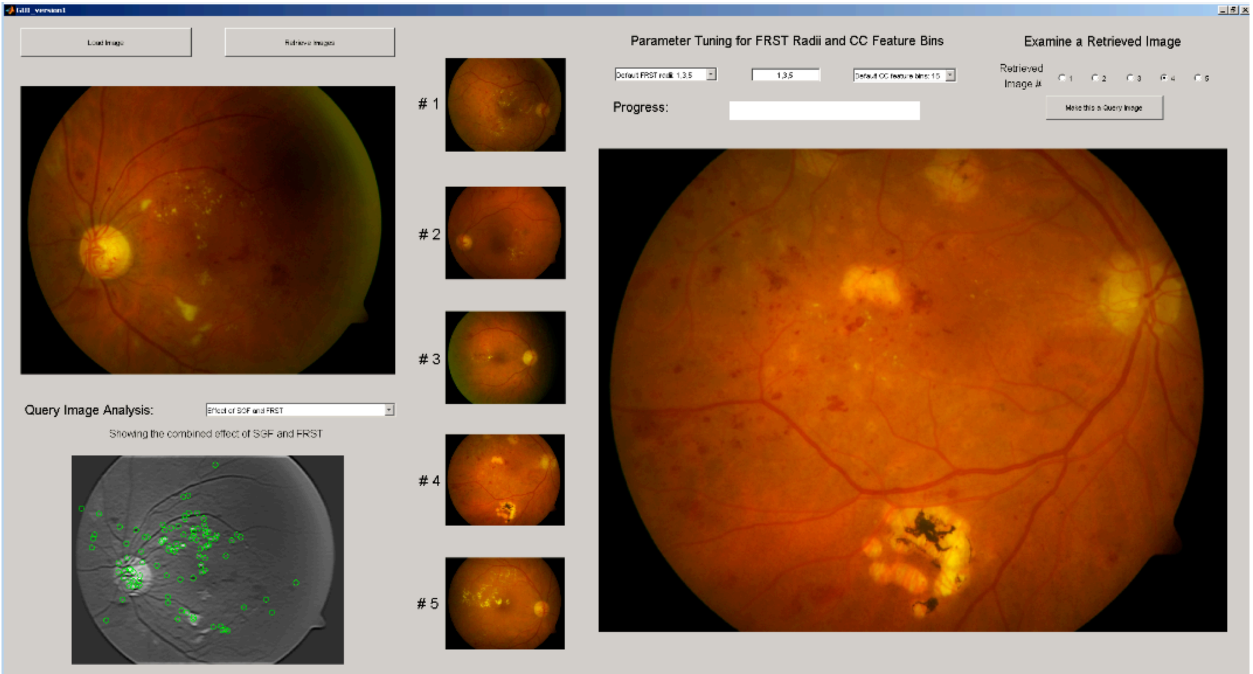

Figure 5. Illustrating the retrieval interface.

List of Publications and Products

Published:

Li, B, Li, H. K., Automated Analysis of Diabetic Retinopathy Images: Principles, Recent Developments, and Emerging Trends. Current Diabetes Reports, 13(4), 453-459, 2013.

Chandakkar P, Venkatesan R, Li B, Retrieving clinically relevant diabetic retinopathy images using a multi-class multiple instance framework, In: Proceedings of SPIE Medical Imaging, Orlando, FL, February 2013.

Chandakkar P, Venkatesan R, Li B, Li H.K., A Machine-learning Approach to Retrieving Diabetic Retinopathy Images, *ACM Conference on Bioinformatics, Computational Biology and Biomedicine* (ACM BCB), 2012.

Venkatesan R, Chandakkar P, Li B, Li H. K., Classification of Diabetic Retinopathy Images Using Multi-Class Multiple-Instance Learning Based on Color Correlogram Features, 34th *International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC), San Diego, CA, August, 2012.

Submitted:

Chandakkar P, Venkatesan R, Li B, MIRank-KNN: Multiple Instance Retrieval of Clinically-Relevant Diabetic Retinopathy Images, submitted to the Journal of Medical Image Analysis.

Software Tools:

A complete software package that can be used to retrieve DR images from a given database, for any given query image, has been built and posted on the following Website:

http://www.public.asu.edu/~bli24/DR-System-and-Data.html

Dataset:

The dataset accompanying the above submitted journal article is also posted on the above Website. The dataset contains the images used in this this project.

Bibliography

[1] Centers for Disease Control and Prevention. 1990. Prevalence and incidence of diabetes mellitus in United States, 1980-1987. MMWR Morb. Mortal. Wkly. Rep. 39:809-812.

[2] Early Treatment and Detection of Diabetic Retinopathy Study Group. 1991. Grading Diabetic Retinopathy from Stereoscopic Color Fundus Photographs -An Extension of the Modified Airlie House Classification. EDTRS Report Nr. 10. Ophthalmology. 98:786-806.

[3] Zana, F., J.-C. Klein, 1999. A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. IEEE Trans. Medical Imaging, 18, 419-428.

[4] Gregson, P.H., Z. Shen, R.C. Scott, V. Kozousek, 1995. Automated grading of venous beading, Computers and Biomedical Research. 28, 291-304.

[5] Pedersen, L., M. Grunkin, B. Ersbøll, K. Madsen, M. Larsen, N. Christoffersen, U. Skands, 2000. Quantitative measurement of changes in retinal vessel diameter in ocular fundus images. Patt. Recog. Lett., 21, 1215-1223.

[6] Gagnon, L., M. Lalonde, M. Beaulieu and M.C. Boucher. 2001. Procedure to Detect Anatomical Structures in Optical Fundus Images. Proceedings of SPIE, 4322, 1218-1225.

[7] Lalonde, M. F. Laliberte, L. Gagnon. 2004. RetsoftPlus: a tool for retinal image analysis Computer-Based Medical Systems. Proceedings of 17th IEEE Symposium on CBMS, 542 – 547.

[8] Natz, N., M. Goldbaum, M. Nelson, and S. Chaudri. 1988. An image processing system for automatic retina diagnosis. *Proceedings of SPIE*, 902, pp. 131–137.

[9] Chaudri, S., S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. 1989. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Trans. Medical Imaging*, 8, pp. 263–269.

[10] Jasiobedzki, P., C.J. Taylor, and J.N. Brunt. 1993. Automated analysis of retinal images, *Image and Vision Computing*, 11, 139-144.

[11] Chaum E, Karnowski TP, Govindasamy VP, Abdelrahman M, Tobin KW., 2008, Automated diagnosis of retinopathy by content-based image retrieval, *Retina,* Vol. 28, No. 10, 1463-77.

[12] Tobin, KW, Abramoff, MD, Chaum, E., Giancardo, L., Govindasamy, VP, Karnowski, TP, Tennant, MTS, Swainson, S., Using a Patient Image Archive to Diagnose Retinopathy, 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, August 20-24, 2008,

[13] Niemeijer M, van Ginneken B, Russell SR, Suttorp-Schulten MS, Abràmoff MD., 2007, Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis, *Invest Ophthalmol Vis Sci.*, Vol. 48, No. 5, pp 2260-7.

[14] Klein R, Klein BE, Moss SE, Wong TY, Sharrett AR., 2006, Retinal vascular caliber in persons with type 2 diabetes: the Wisconsin Epidemiological Study of Diabetic Retinopathy: XX., Ophthalmology. 2006 Sep;113(9):1488-98. Epub 2006 Jul 7.

[15] Xu, X., Li, B., 2008, Automatic Classification and Detection of Clinically-Relevant Images for Diabetic Retinopathy, *SPIE Medical Imaging*, vol. 6915 (2), pp. 69150Q.1-69150Q.9.

[16] Ravishankar, S., Jain A., Mittal, A., 2009, Automatic feature extraction of early detection of diabetic retinopathy in fundus images, Proc. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 210-217.

[17] Klein, R. and BEK. Klein. 1995. Vision disorders in diabetes. pp.293-339. *In*: R. Klein, BEK. Klein (ed.), Diabetes in America,2nd ed. Bethesda, Md: National Institutes of Health and National Institute of Diabetes and Digestive and Kidney Diseases.

[18] Amos, A., D. McCarty and P. Zimmet. 1997. The rising global burden of diabetes and its complications: estimates and projections to the year 2010. Diabet. Med. 14, S7–S13.

[19] Zhang X., Saaddine, J. B., Chou C., Cotch, M., Cheng, Y., Geiss, L. S., Gregg, E. G., Albright, A. L., Klein, B. E., Klein, R., 2010, Prevalence of Diabetic Retinopathy in the United States, 2005-2008, The Journal of the American Medical Association, 2010 August: 304(6): 649-656.

[20] B. Manjunath and W. Ma, "Texture features for browsing and retrieval of image data," *IEEE transactions on pattern analysis and machine intelligence (PAMI-special issue on digital libraries),* vol. 18, pp. 837-842, August 1996.

[21] Q. Chen, X. Tai, Y. Dong, S. Pan, X. Wang and C. Yin, "Medical image retrieval based on sementic of neighborhood color moment histogram," in *The 2nd international conference on bioinformatics and biomedical engineering (ICBBE)*, 2008.

[22] Chang, Chih-Chung and C.-J. Lin, "{LIBSVM}: A library for support vector machines," *ACM transactions on intelligent systems and technology,* vol. 2, no. 3, pp. 27:1 - 27:27, 2011.

[23] M. Li, "Texture moment content-based image retrieval," in *IEEE ICME*, 2007.